LA-UR-18-24634
Approved for public release; distribution is unlimited.

Title: A thought experiment to evaluate a proposed data analysis method

Author(s): Lestone, John Paul

Intended for: Correspondence with the DOE NDSE review committee

Issued: 2018-05-30

A thought experiment to evaluate a proposed data analysis method

J. P. Lestone
May 17th, 2018
Computational Physics Division
Los Alamos National Laboratory

Abstract

A statistical analysis was recently presented at a DOE review in reference to a particular problem. Here we attempt to capture the essence of this problem with a simple "toy" analog and demonstrate the incorrectness of the presented *J* figure of merit algorithm. We show that the incorrect usage of *J* gives PDFs that have erroneous long tails and significantly overestimate the rms deviations relative to those obtained by exact analytic methods. Conversely, a numerical Bayesian analysis gives the correct result.

(1) Introduction

Here we study a very simple data analysis problem where a detector observes a signal $y_{A1}\pm\Delta y_{A1}$ at time t_1, and $y_{A2}\pm\Delta y_{A2}$ at time t_2 in experiment A; and $y_{B1}\pm\Delta y_{B1}$ at time t_1 and $y_{B2}\pm\Delta y_{B2}$ at time t_2 in experiment B. We assume the data uncertainties are Gaussian and that the signals vary linearly with time: $y=C+Mt$. M can be thought to be a surrogate for die-away, and C related to timing uncertainties [1,2], for an NDSE application described in more detail in [3]. The question to be asked is: what are the slopes M for experiments A and B? For our simple "toy" problem the slope from experiment A is known analytically,

$$M_A = \frac{y_{A2}-y_{A1}\pm\sqrt{\Delta y_{A1}^2+\Delta y_{A2}^2}}{t_2-t_1} \tag{1}$$

and similarly for experiment B. We imagine the experimental results y_{A1}=2.00±0.02, y_{A2}=1.00±0.01, y_{B1}=2.00±0.02, y_{B2}=0.600±0.006 with t_1=0.00 and t_2=1.00 (see Fig. 1). The corresponding measured slopes are known analytically to be M_A=–1.000±0.022 and M_B=–1.400±0.021 with Gaussian uncertainties. The two slopes are clearly distinguishable. Below we apply the algorithm used by Lawrence Livermore National Laboratory (LLNL) [1,2] to the simple test problem outlined above to point out some of its deficiencies. This is followed by a parallel numerical analysis using standard Bayesian inference that gives the same outcomes as the known analytic result as given by Eq. (1).

References
[1] M. R. Zika *et al*., LLNL report COPD-2018-0089
[2] M. R. Zika, LLNL report COPD-2018-0119.
[3] J. P. Lestone and C. R. Bates, XCP-3:18-003-C, LA-CP-pending.

(2) Zika *et al*. *J* FOM algorithm [1,2]

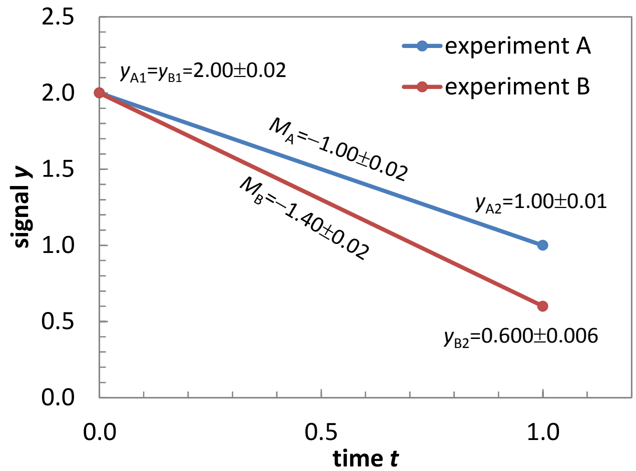

Fig. 1. Experimental results for the toy problem outlined in section (1).

We here use the figure of merit (FOM) [1]

$$J = \frac{1}{\Delta t}\int_t^{t+\Delta t} \frac{\left|I_\gamma^{ref}(t) - I_\gamma(t)\right|}{I_\gamma^{ref}(t)}\,dt.$$

In the case of our simplified problem with only two data per experiment this translates to

$$J = \frac{\left|\frac{y_1^{ref}-y_1}{y_1^{ref}}\right| + \left|\frac{y_2^{ref}-y_2}{y_2^{ref}}\right|}{2}.$$

If we make experiment A the reference then the FOM for experiment B is J_B=0.20.

We construct a series of 10000 trial models where the intercept C and slope M are chosen randomly. Here we assume C is from a Gaussian distribution with a mean of 2.0 and a standard deviation of 0.1, and M is uniformly distributed from 0 to −2. The first 200 of the 10000 trial models are displayed in Fig. 2. For each model we calculate the FOM J_i with i=1 to 10000 for each of the trial models. The array of corresponding J_i versus (−1)×M_i is displayed in Fig. 3. The red-dashed lines show a slice of J, 0.01 either side of J_B=0.20. Fig. 4 shows the relative probability density function (PDF) obtained via the points in the narrow slice about J_B=0.20. This PDF is not the uncertainty PDF for the quantity M for experiment B. However, LLNL assumed it was. Please remember that from section (1) we know the slope from experiment B is M_B=−1.400±0.021. Please note the PDF in Fig. 4 is bimodal, with another peak in the PDF that is obviously unphysical. This could be masked by only including trial models with a slope less than that of experiment A (as done by Zika *et al.*) giving the PDF as shown to the right of (−1)×M=1.00 in Fig. 4. Please notice the non-Gaussian nature of the PDF to the right of (−1)×M=1.00. Using only the right hand side of the PDF gives M=−1.37±0.10 (here we quote the rms spread as a measure of the uncertainty). The rms uncertainty is ~5 times larger than the correct analytical result of 0.021. The Zika et al J metric inferred PDFs are very sensitive to the distribution of trial models (the priors). The width, "peakyness", and skewness of the PDF in Fig. 4 can be modified by changes in the priors. This sensitivity to the priors has little to do with the "real" uncertainty for M_B.

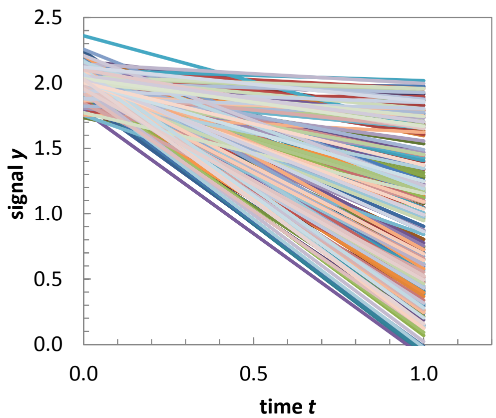

Fig. 2. The first 200 of the 10000 trial models.

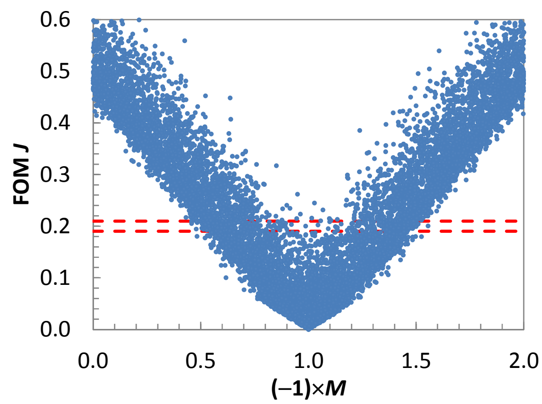

Fig. 3. J versus $(-1)\times M$ for the 10000 trial models discussed in the text. Here the J values are calculated using experiment A as the reference (as done by Zika *et al.*). Notice the FOM is small for models close to experiment A (with a slope M=−1.00) and grows larger as the trial model "slope" moves away from M = −1.00. The red-dashed lines show a slice of J, 0.01 either side of J_B=0.20.

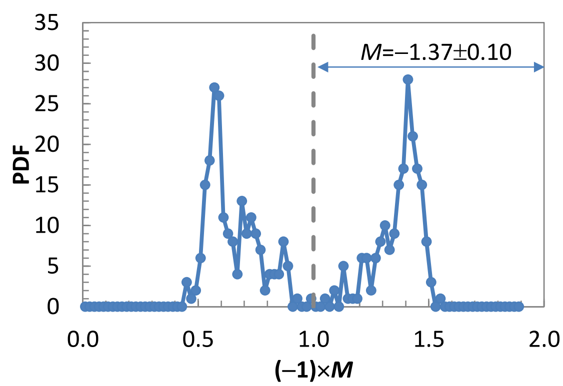

Fig. 4. Relative PDF obtained via the points in the narrow slice about J_B=0.20 shown in Fig. 3. This is not the uncertainty PDF for the quantity M for experiment B. Using only the right hand side of the PDF gives M=−1.37±0.10. The rms uncertainty of 0.10 is ~5 times larger than the correct analytical result of 0.021.

Now consider the possibility that "other" information is available that reduces the uncertainty in the intercept C in our thought experiment by a factor of four. With this change, the results corresponding to figures 3 and 4 are displayed in figures 5 and 6. Notice that the PDF on the right hand side of Fig. 6 now resembles the known analytic result. However, it is not the uncertainty PDF for the quantity M for experiment B.

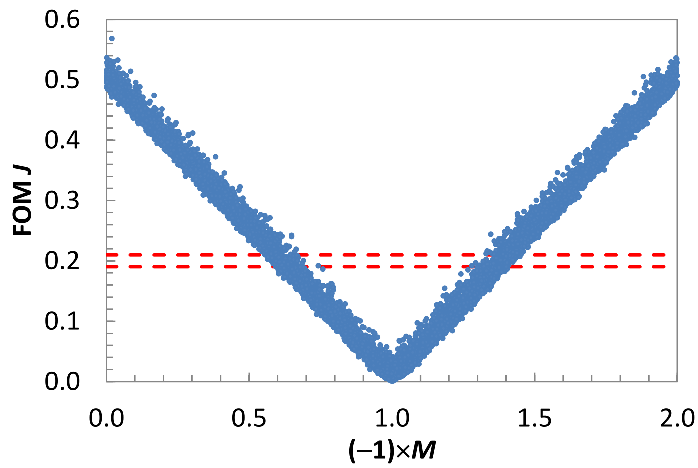

Fig. 5. As for Fig. 3 but with the uncertainty in the intercept C reduced to 0.025 (1-σ).

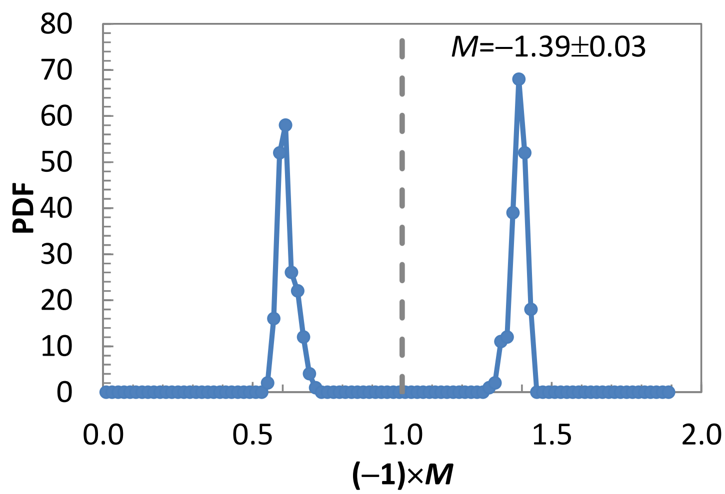

Fig. 6. Relative PDF obtained via the points in the narrow slice about J_B=0.20 shown in Fig. 5. This is not the uncertainty PDF for the quantity M for experiment B.

Please remember that for our toy problem we know the slope for experiment B is M_B=−1.400±0.021. The algorithm choice of Zika *et al*. gives the illusion that if C is not constrained in the prior then the uncertainty in M will be large and non-Gaussian (see Fig. 4), while if C is constrained in the prior then the uncertainty in M is small (see Fig. 6). The root cause of the problem is the metric J has a useful meaning only when it is small (about or smaller than the relative experimental uncertainties). Locking in the expectation of experiment A as the "reference" and using the large FOM J_B obtained with the expectation of experiment B relative to the reference A, to generate the uncertainty PDF for experiment B, is inappropriate.

(3) Standard numerical Bayesian solution for the "toy" problem

First we calculate the χ^2 using the data from experiment A relative to each of the 10000 trial models. The corresponding array of χ^2 versus M points is displayed in Fig. 7. Each point is given the weight $W=\exp(-\chi^2/2)$. Projecting these weights down onto the horizontal axis gives the uncertainty PDF for M_A and is displayed in Fig. 8. The corresponding results for experiment B are displayed in figures 9 and 10. It would be inappropriate to use the χ^2 difference between experiment A and B of ~1600, and then obtain an uncertainty PDF for experiment B by taking a slice of the array elements from Fig. 7 about χ^2~1600. This would give a mean value near $M_B=-1.4$ (only using the right hand side of the figure) but the distribution would be wide with long non-Gaussian tails.

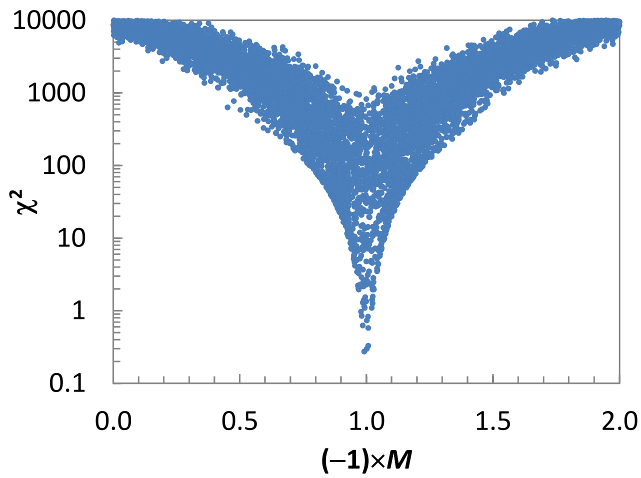

Fig. 7. χ^2 versus M points for experiment A.

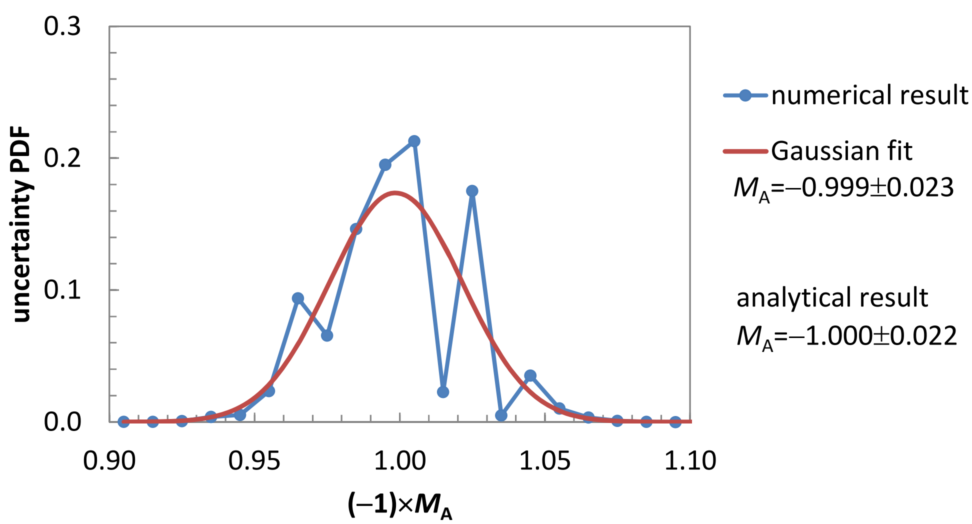

Fig. 8. Slope M_A uncertainty PDF obtained by brute-force numerical Bayesian inference. This numerical result differs from a Gaussian, only because of the small number of trials that are a close match to the data. This can be rectified by resampling the 10000 trials over a space more conducive to the size of the true uncertainties.

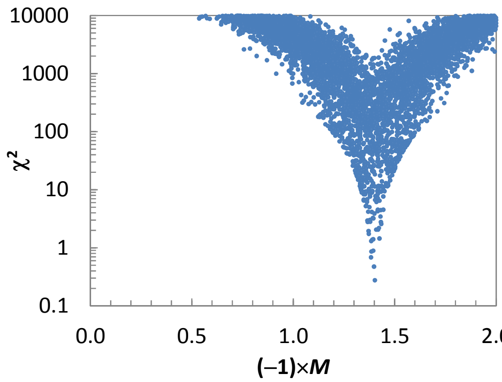

Fig. 9. χ^2 versus M points for experiment B.

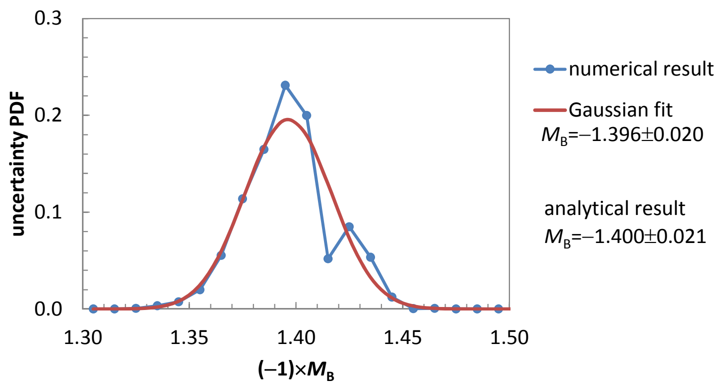

Fig. 10. Slope M_B uncertainty PDF obtained by brute-force numerical Bayesian inference. This numerical result differs from a Gaussian, only because of the small number of trials that are a close match to the data. This can be rectified by resampling the 10000 trials over a space more conducive to the size of the true uncertainties, as presented in Fig. 11.

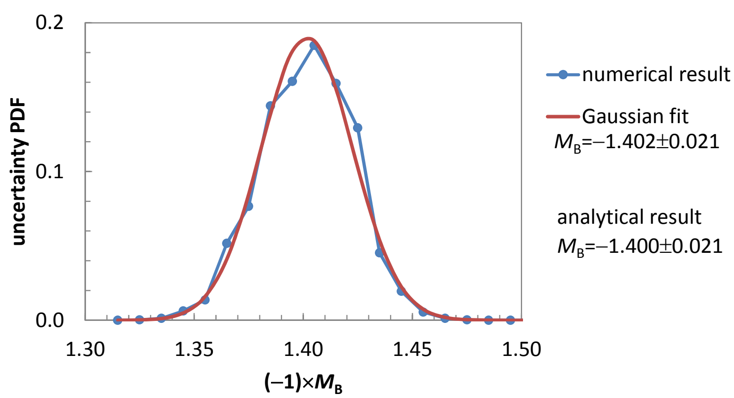

Fig. 11. As for Fig. 10, but with a more efficient prior set. Here M is sampled uniformly from −1.30 to −1.50, instead of from 0 to −2.0 as done for Fig. 10.

Please notice the brute-force numerical Bayesian inference is in agreement with the slopes obtained using Eq. (1) reported in section (1), while the Zika *et al*. method generates incorrect results that are very sensitive to the assumed priors. When broad priors are used in a 2 parameter model the method of Zika *et al*. generates non-Gaussian distributions with non-physical long tails, and rms uncertainty values much larger than the corresponding known analytical solutions.